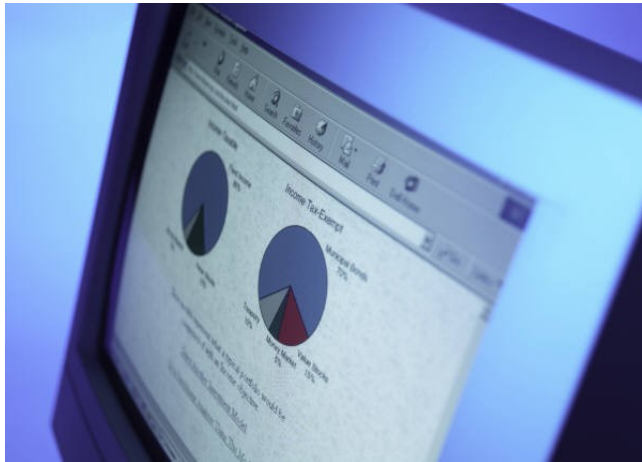


DATA VALIDATION & READINESS



**Transforming data into information,
one byte at a time.**



Finance & Administration Cabinet
101 Cold Harbor
Frankfort, KY 40601

<http://technology.ky.gov>
(502)564-7576

The Benefits

- Metadata captured.
- Improved efficiencies.
- Increased level of trust in data, results, reports, etc.
- Business model changes handled more efficiently, effectively, and timely.
- Correct and informed decisions made the first time.
- Increased customer satisfaction.
- Legal liabilities reduced.
- Eliminate unnecessary processes which are in place because the quality of data is questionable.
- Eliminate/reduce operational system interferences and down-time.
- Projects involving application rework, merging systems, or migrating data to a new system completed on schedule and/or ahead of schedule because data is of good quality and the metadata is known.
- Provide business intelligence on individuals and organizations for research, fraud detection and planning.



Who We Are

The Data Services Branch was formed in 2006 to address enterprise data needs within the Commonwealth of Kentucky.

The team works with our customers to profile and analyze their data to expose data anomalies, apply business rules to standardize, cleanse, match & merge as well as transform their data into quality data.

If Data Quality is important to you we can:

- Access your data, wherever it may be.
- Understand your data and how reliable it is.
- Cleanse, standardize, validate and correct your data.
- Expose duplicate data.
- Match and/or merge your data.
- Transform data so it can be understood by applications, people and processes that need it.
- Deliver data to new destinations for new uses
- Geocode your data to include latitude and longitude in order to provide maps when applicable.

Data Services Branch

Our Tools

- Information Integrator Classic Federation
- Information Integrator Federated Server
- Information Analyzer
- QualityStage
- DataStage
- MetaData Server Repository
- Business Glossary



Data Integration & Data Conversion

In the current economy with everyone looking to save money Data Integration of systems can reduce overhead and ease redundancy of data. It can enable the sharing of data within an agency or among agencies for reporting, strategic planning, business intelligence, and forecasting just to name a few.

With the Commonwealth looking to replace legacy systems with newly developed systems or commercial off-the-shelf systems (COTS), data conversion is a necessity. The DSB has the tools and experience to access the data where it sits in the legacy system, use our tools and process to understand that data, work with the business to gather business rules and apply those rules to convert the data and move it to the new system or to a format which the new system can pick up.

Data Standardization, Cleansing & Transformation

Standardization may include standardizing addresses to match the United States Post Office format, or standardizing date or phone number columns/fields so that they meet the required format.

Cleansing may include parsing data such as a name field so that it is broken down to first, middle, last rather than all included in one column/field or making sure columns/fields that should be numeric do not contain alphas.

Transformations may include moving certain types of data from one column/field to another, parsing that information out, or doing some aggregations on a set of columns/fields.

Data Profiling & Analysis

Profiling and analysis is a process of discovery and investigation of the current data to expose and understand the current quality of the data and to highlight the data issues and anomalies.

While profiling we look at the structure of the data, the definitions of the data and the content of the data which allows for comprehensive understanding of the quality of the current data.

We are finding anomalies such as;

- data within the system does not match the definition of what should be there.
- data spread across columns/fields that should be contained in one.
- dates and/or numbers that are out of range for what should be there.
- columns/fields that should contain numbers only but also contains alphas.

As we complete the profiling and analysis a report is created to highlight our findings and provide samples of the anomalies.

